Volume 12 Number 10 1984 Nucleic Acids Research

The amino-acid sequence of two non-toxic mutants of diphtheria toxin: CRM45 and CRM197

G.Giannini, R.Rappuoli and G.Ratti*

Sclavo Research Centre, Via Fiorentina, 1, 53100 Siena, Italy

Received 13 April 1984; Accepted 2 May 1984

SUMMARY

The amino-acid sequences of two diphtheria toxin-related, non-toxic proteins, CRM45 and CRM197, were deduced from the complete sequence of their genes: tox 45 and tox 197. CRM45 lacks the last 149 C-terminal amino-acid residues, but is otherwise identical to diphtheria toxin: a single C→T transition introduces an "ochre" (TAA) termination signal in tox 45, after the codon for threonine-386. A single G→A transition was also found in tox 197, leading to the substitution of glycine-52, present in the wild-type toxin, with glutamic acid in CRM197. This aminoacid change is responsible for the loss of the NAD:EF2 ADP-ribosyltransferase activity in CRM197, due most probably to an alteration of the NAD^+ binding site.

INTRODUCTION

Uchida et al. (1,2,3) described in 1973, five diphtheria toxin-related proteins, obtained by mutation with nitrosoguanidine of corynephage ß DNA containing the gene tox for diphtheria toxin. Following infection and lysogenization of Corynebacterium diphtheriae, a number of mutated tox genes were expressed by the host bacterium and purified from culture supernatants. These products were given the general name of CRM (cross reacting materials). Two of these, CRM45 and CRM197, have attracted in the last few years the attention of laboratories interested in studying the properties of chimeric toxins, e.g. as specific cytocidal agents (4), or new vaccine development (5,6,7).

Diphtheria toxin kills sensitive eukaryotic cells by: i) binding to cell surface receptors; ii) entering the cytosol via endocytosis and membrane translocation, mediated by the low pH of the endocytic compartment; iii) blocking protein synthesis by inactivation of the ribosomal elongation factor EF2. This last event is caused by the enzymatic activity (NAD:EF2 ADP-ribo-

syltransferase) possessed by the 193 aminoacid long, N-terminal peptide, known as fragment A. (For review see ref. 8 and 9). The fact that different functions required for the cell intoxication process can be assigned to different segments of the molecule, makes diphtheria toxin, and its functional mutants, suitable for building composite proteins, retaining only some of the original properties of the toxin.

Both CRM45 and CRM197 have been well characterized (9) and found to be non-toxic: CRM45 was shown to lack a C-terminal peptide, approx. 17,000 daltons, and as a consequence, to be incapable of binding to the cell surface, whereas CRM197's lack of toxicity was shown to be due to the loss of enzymatic activity of its fragment A.

We now report the primary structure of these two proteins and the precise extent and nature of the mutation caused by the nitrosoguanidine treatment.

METHODS

Lysogenic C. diphtheriae strains C7(ß197) and C7(ß45), a gift from A.M. Pappenheimer Jr., were the source of ß45 and ß197 corynephages. These phages were used to infect C. diphtheriae C7(-) strain which was then grown in 10 l fermenters, as described (10). ß45 and ß197 phages were then purified from bacterial lysates (10) and the DNA was extracted and purified according to standard methods (11). Phage DNA was digested with restriction endonuclease BamHI and the Bam-4 (3.9 Kb) fragment, previously shown to contain the gene tox in ß corynephages (12), was purified by agarose gel electrophoresis. A 1.88 Kb fragment comprising the entire tox operon (13) was then obtained by digestion of Bam-4 with EcoRI + HindIII. The Eco-Hind fragment was further digested with a set of restriction endonucleases as previously described for the cloning and sequencing of the wild-type omega tox gene (13).

Sequencing was performed according to Sanger et al. (14) on recombinant M13mp8 and mp9 (15) single-stranded templates.

RESULTS

The two 1.88 Kb HindIII-EcoRI restriction fragments, containing tox45 and tox197 genes, were purified from ß45 and ß197 corynephage DNA.

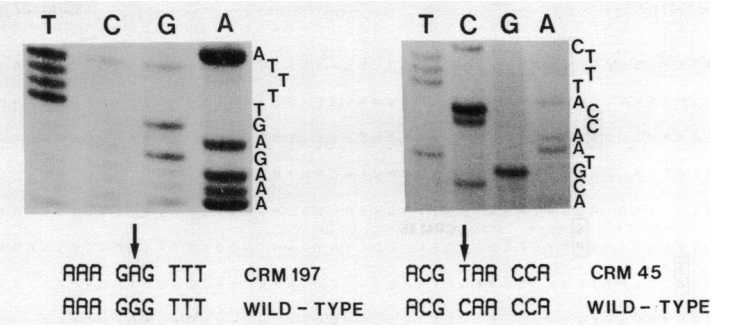

Figure 1. Portions of two sequencing gels showing the mutated nucleotides in tox45 (right) and tox197 (left). The two mutated sequences are also shown against the corresponding wild-type version (16); the spacing indicates the reading frame.

Their complete nucleotide sequence was determined and compared with that of the corresponding wild-type fragment in corynephage ß (16).

Only one nucleotide change was found in either mutant sequence. In both cases a (C to T;G to A) transition had occurred: at 1355 b.p. from the first base of the HindIII site for tox45, and 351 b.p.for tox197.

The sequences at the two mutation sites as present in the "sense" strand are shown in Figure 1, compared with the wild-type version. The amino-acid sequences of CRM45 and CRM197 were deduced from the corresponding genes and compared with wild-type diphtheria toxin. The results are shown in Figure 2. The mutation in tox45 changes the wild-type "CAA" codon for glutamine-387 into the termination signal "TAA", whereas in tox197 the wild-type codon "GGG" for glycine is mutated into "GAG" for glutamic acid at pos. 52.

DISCUSSION

The nucleotide sequences of the two 1880 b.p.-long DNA segments containing the genes tox45 and tox197 showed that in both cases the mutagenic treatment with nitrosoguanidine performed by Uchida et al. (1) caused only one point mutation in the entire segment. In both cases the single mutation occurred within the coding portion and consisted of a G:C to A:T transition, in accordance with the predominant mode of action of nitrosoguanidine (17).

Figure 2. The amino-acid sequence of wild-type diphtheria toxin is shown in one-letter symbols. The vertical broken line at pos. 193 indicates the C-terminal of fragment A. The alterations in the primary structure for the three mutants up to now described in detail, are shown above the main sequence. CRM197: framed residue at pos. 52. CRM45: C-terminal residue marked by asterisk. CRM228: shadowed residues (deduced from Kaczorek et al., ref. 18).
(A = alanine; C = cysteine; D = aspartic acid; E = glutamic acid; F = phenylalanine; G = glycine; H = histidine; I = isoleucine; K = lysine; L = leucine; M = methionine; N = asparagine; P = proline; Q = glutamine; R = arginine; S = serine; T = threonine; V = valine; W = tryptophan; Y = tyrosine).

This was somewhat unexpected since the sequence of a similarly mutated gene, tox288, shows a much higher degree of change. Gene tox228 was produced by Uchida et al. at the same time (1) as tox45 and tox197 and its primary structure, recently determined by Kaczorek et al. (18) shows when compared with the wild-type tox gene, 16 nucleotide changes within the coding portion of the gene, five of which cause amino-acid changes in the gene product, CRM228 (see Fig. 2).

CRM45

Our data essentially confirm the early characterization of CRM45 by Uchida et al. (19): the "TAA" termination signal introduced by the (C to T) change in the "CAA" codon for glutamine-387 causes early termination at threonine-386 in CRM45 which therefore lacks the C-terminal last 149 amino-acid residues (16,530 daltons). The deduced molecular weight of CRM45 is therefore 41,826 daltons, in good agreement with SDS-polyacrylamide gel electrophoresis data (M_r=41-45,000). Bacha & Murphy (20) described two C. diphtheriae nonsense suppressor strains, Sup 1 and Sup 2, capable of

translating the tox45 gene into a full-length product, enzymatically and immunologically identical to diphtheria toxin. The data here presented indicate that these strains carry an "ochre" suppressor mutation.

It is interesting to note that in the case of another nonsense mutant gene, tox30, normally coding for a polypeptide of approximately 30,000 daltons, the mutation is suppressed by Sup 1 strain but not by Sup 2. Since we now show that both Sup 1 and Sup 2 suppress the "ochre" signal, the best explanation for the behaviour of tox30 in these strains is that this gene probably carries an "amber" mutation which is suppressed by Sup 1 ("ochre/amber" suppressor in accordance with the behaviour of E. coli "ochre" suppressors, all of which can recognize "amber" non sense mutations as well) but not by Sup 2. This second strain would then be an "ochre"-only suppressor incapable of suppressing the "amber" mutation, as described in yeast (21). If this interpretation of Bacha & Murphy's results were correct, in the tox30 gene a N-terminal polypeptide of approx. 30,000 daltons could be produced most likely by a (G to A) transition, converting the "TGG" codon of tryptophan-281 into "TAG".

CRM197

The single point mutation found in tox197 alters one amino-acid residue in fragment A (see Fig. 2). The fragment has the features of a functional domain and can be released from the whole toxin molecule by mild proteolysis and reduction: this process unmasks the enzymatic activity of diphtheria toxin, NAD:EF2 ADPribosyl-transferase, which was shown to reside entirely within fragment A itself (9). The transferase activity has been extensively characterized (9): fragment A binds first NAD^+ at a single binding site (22) and subsequently a portion of the eukaryotic ribosomal factor EF2 comprising a modified histidine (diphthamide); eventually, the transfer of an ADP-ribosyl group onto the diphthamide residue of EF2 yields nicotinamide and the inactive ADP-ribosyl-EF2.

CRM197 was shown to possess an enzymatically inactive fragment A. The observation that intact wild-type diphtheria toxin can bind NAD^+ whereas CRM197 appears unable to do so (23), indicated that some alteration of the NAD^+ binding site in CRM197 is the primary, if not the only, cause of loss of

enzymatic activity, and thereby, toxicity. The present finding that the substitution of glycine-52 with glutamic acid is the only difference between CRM197 and diphtheria toxin strongly suggests that glycine-52 plays an essential role in the NAD^+ binding site of diphtheria toxin. This is supported by other information available about the structure of fragment A's active sites. Fluorescence quencing studies (22) give evidence for the possible formation of a charge transfer complex between the nicotinamide moiety of NAD^+ and at least one of the two tryptophan residues of fragment A. On the other hand, it has been shown (24) that tryptophan-153 is either involved in EF2 binding or is part of the catalytic site, but is not involved in NAD^+ binding: this points to tryptophan-50 as being part of the NAD^+ binding fold. The mutation of CRM197 at pos. 52 is therefore in agreement with these observations. The involvement of one essential tyrosine residue in the binding of NAD^+ has also been reported (25). Although a specific tyrosine residue could not be singled out, it is interesting that in the segment between tyrosine-20 and tyrosine-85 are present 7 out of the 10 tyrosine residues in fragment A.

Figure 2 also shows five amino-acid mutations present in another mutant toxin, CRM228, the structure of which was recently described by Kaczorek et al. (18), which, like CRM197, possesses an inactive fragment A. There are two mutations in the fragment A moiety of CRM228 and, interestingly, both involve, as for CRM197, a charge alteration. To our knowledge, experimental data on NAD^+ binding to CRM228 are not available and the loss of enzymatic activity could be attributed to either of the two changes. Nevertheless, the substitutions in CRM228 of another glycine residue in the same portion of the molecule which is mutated in CRM197, is suggestive since glycine residues have been shown to be strongly conserved in "consensus" sequences thought to be general features of adenosine nucleotide binding folds in protein of various origin (26).

ACKNOWLEDGEMENTS

We would like to thank G. Corsi for the graphic work and Susan Weemys Taddeo for preparing the typescript.

*To whom correspondence should be sent

REFERENCES

1. Uchida, T., Pappenheimer, A.M. Jr. and Greany, R. (1973) J. Biol. Chem. 248, 3838-3844.
2. Uchida, T., Pappenheimer, A.M. Jr. and Harper, A.A. (1973) J. Biol. Chem. 248, 3845-3850.
3. Uchida, T., Pappenheimer, A.M. Jr. and Harper, A.A. (1973) J. Biol. Chem. 248, 3851-3854.
4. Bacha, P., Murphy, J.R. and Reichlin, S. (1983) J. Biol. Chem. 258, 1565-1570.
5. Porro, M., Saletti, M., Nencioni, L., Tagliaferri, L. and Marsili, I. (1980) J. Infect. Dis. 142, 716-724.
6. Anderson, P. (1983) Infect. Immun. 39, 233-238.
7. Rappuoli, R. (1983) Appl. Environ. Microbiol. 46, 560-564.
8. Pappenheimer, A.M. Jr. (1977) Ann. Rev. Biochem. 46, 69-94.
9. Collier, R.J. (1982) In ADP-ribosylation Reactions, Hayaishi, D. and Ueda, K. Eds., pp. 575-592, Academic Press, New York.
10. Rappuoli, R., Michel, J.L. and Murphy, J.R. (1983) J. Virol. 45, 524-530.
11. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1983) Molecular Cloning - A Laboratory Manual. Cold Spring Harbor Laboratory, CSH, New York.
12. Costa, J., Michel, J., Rappuoli, R. and Murphy, J.R. (1981) J. Bacteriol. 135, 511-516.
13. Ratti, G., Rappuoli, R. and Giannini, G. (1983) Nucleic Acids Res. 11, 6589-6595.
14. Sanger, F., Coulson, A.R., Hong, G.F., Hill, D.R. and Petersen, G.B. (1982) J. Mol. Biol. 162, 729-773.
15. Messing, J. and Vieira, J. (1982) Gene 19, 269-276.
16. Greenfield, L., Bjorn, M.J., Horn, G., Fong, D., Buck, G.A., Collier, R.J. and Kaplan, D.A. (1983) Proc. Natl. Acad. Sci. USA 80, 6853-6857.
17. Miller, J.H. (1972) in Experiments in Molecular Genetics, Cold Spring Harbor Laboratory, CSH New York, p. 117.
18. Kaczorek, M., Delpeyroux, F., Chenciner, N., Streeck, R.E., Murphy, J.R., Boquet, P. and Tiollais, P. (1983) Science 221, 855-858.
19. Uchida, T., Gill, M.D. and Pappenheimer, A.M. Jr. (1971) Nature New Biol. 233, 8-11.
20. Bacha, P. and Murphy, J.R. (1978) J. Bacteriol. 136, 1135-1142.
21. Gilmore, R.A., Stewart, J.W. and Sharman, F. (1971) J. Mol. Biol. 61, 157-174.
22. Kandel, J., Collier, R.J. and Chung, D.W. (1974) J. Biol. Chem. 249, 2088-2097.
23. Lory, S., Carroll, S.F. and Collier, R.J. (1980) J. Biol. Chem. 256, 12016-12019.
24. Michel, A. and Dirkx, J. (1977) Biochim. Biophys. Acta 491, 286-295.
25. Beugnier, N. and Zanen, J. (1977) Biochim. Biophys. Acta 490, 225-234.
26. Walker, J.E., Saraste, M., Runswick, M.J. and Gay, N.J. (1982) EMBO Journal 1, 945-951.